Technical Memo

ANL/EES-TM-301

A REGIONALIZATION METHODOLOGY FOR SECTOR MODEL INPUT DATA: DERIVATION AND APPLICATIONS

ARGONNE NATIONAL LABORATORY
Energy and Environmental Systems Division

Operated by
THE UNIVERSITY OF CHICAGO for U. S. DEPARTMENT OF ENERGY
under Contract W-31-109-Eng-38

Argonne National Laboratory, with facilities in the states of Illinois and Idaho, is owned by the United States government, and operated by The University of Chicago under the provisions of a contract with the Department of Energy.

This informal report presents preliminary results of ongoing work or work that is more limited in scope and depth than that described in formal reports issued by the Energy and Environmental Systems Division.

Printed in the United States of America. Available from National Technical Information Service, U. S. Department of Commerce, 5285 Port Royal Road, Springfield, Virginia 22161.

ARGONNE NATIONAL LABORATORY
9700 South Cass Avenue, Argonne, Illinois 60439

ANL/EES-TM-301

A REGIONALIZATION METHODOLOGY FOR SECTOR MODEL INPUT DATA: DERIVATION AND APPLICATIONS

by

D.A. Hanson, D.W. South, and W.H. Oakland*

Energy and Environmental Systems Division
Policy and Economic Analysis Group

June 1985

(revised November 1988)

work sponsored by

U.S. DEPARTMENT OF ENERGY
Assistant Secretary for Fossil Energy
Office of Planning and Environment

*Department of Economics, Tulane University.

CONTENTS

FIGURES

FOREWORD

Under the auspices of the National Acid Precipitation Assessment Program (NAPAP), activities supporting the preparation of future assessments have been planned and delegated to task groups. Task Group B (TG-B), "Man-Made Sources" (subsequently redesignated Task Group I, "Emissions and Controls"), of the Interagency Task Force on Acid Precipitation is responsible for developing and testing models that can be used to project fuel use and air pollutant emissions by sector. Argonne has participated in the TG-B program since 1984.

The TG-B program is being carried out in two phases. Phase 1 includes development of the models for generation of baseline scenarios. Phase 2 will address the capabilities for modeling emission control scenarios. Under Phase 1, the sector models are being developed and tested. This testing is designed to aid in model development and help prepare the models for use by the task force. Upon completion, the sector models will be incorporated into the TG-B emissions model set and linked to a system of models that provide scenario-consistent input data.

Within the Policy and Economic Analysis Group of Argonne's Energy and Environmental Systems Division, the Energy-Economic Modeling Program is publishing a series of reports that document the steps undertaken to prepare national and regional projections of energy and economic activity required as input to the sector emissions models. This report is part of this series; it documents the methodology used to translate national control forecasts into the specific regional data needed to drive the sector models. Separate reports are being prepared for each sector model because the driver data are highly specific.

Although the driver data for each sector model are different in configuration, a common regionalization scheme is used. The Argonne Regionalization Activity Module (ARAM) was developed to systematically generate regional forecasts of energy and economic variables required by the sector models in the TG-B emissions model set. This report focuses on the derivation and interpretation of the regionalization formula incorporated in ARAM.

A REGIONALIZATION METHODOLOGY FOR SECTOR MODEL INPUT DATA: DERIVATION AND APPLICATIONS

by

D.A. Hanson, D.W. South, and W.H. Oakland

1 INTRODUCTION

In the National Acid Precipitation Assessment Program (NAPAP), Task Group B (TG-B) is responsible for developing and testing models that can be used to project fuel use and air pollutant emissions by energy use sector. As discussed in the Foreword, this work is being carried out in two phases. All activities described in this report have taken place under Phase 1 of the TG-B program. This report addresses one aspect of the system designed to supply energy-economic driver data to the TG-B emissions model set: the general regionalization methodology employed. The components of the energy-economic driver module are shown in Fig. 1.

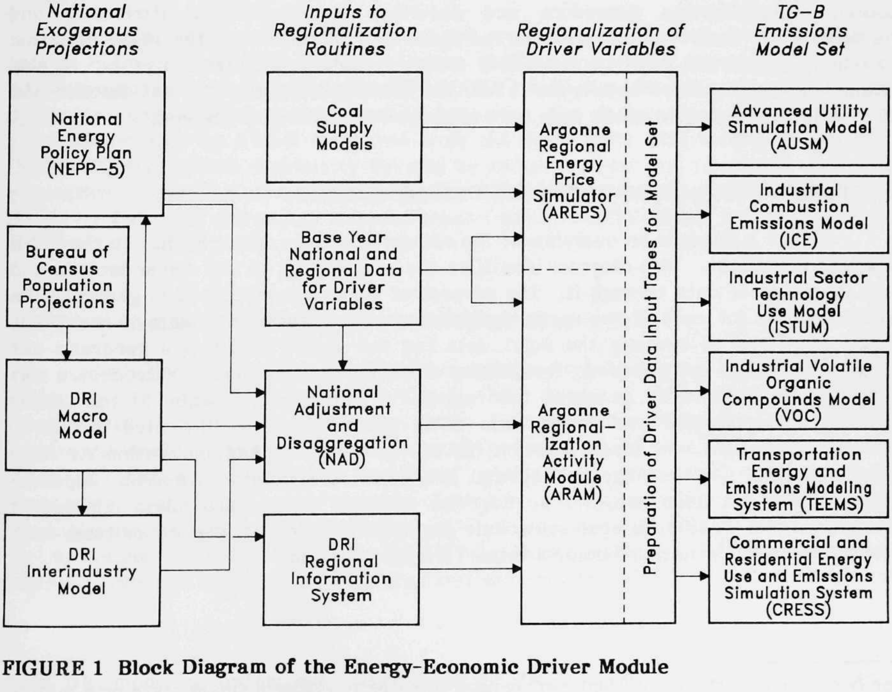

FIGURE 1 Block Diagram of the Energy-Economic Driver Module

1.1 BACKGROUND AND SCOPE

As indicated, this report describes the general approach used to regionalize energy (quantity) and economic data required by sector models in the TG-B emissions model set. The TG-B model set is a collection of sector emissions models, each of which operates at a regional level.* These models project the emissions of the precursors of acid precipitation and related energy use. Each of these models must be provided with energy or economic input variables at the regional level.

To supply these data, a method was devised to generate regional forecasts of the energy and economic variables required by these sector models. The regionalization method, which is based on regional economic theory, is incorporated in the Argonne Regionalization Activity Module (ARAM), which is a component of the model set (see Fig. 1). ARAM systematically transforms the control values produced by national forecasting models into the regional projections required as input by the sector models. This report describes the regionalization approach incorporated in ARAM and the TG-B emissions model set.

The discussion of the ARAM methodology is organized as follows. The remainder of this introduction presents background on the need for and approach to regionalizing energy and economic projections in the TG-B emissions model set. Section 2 presents the adopted regionalization procedure and describes its properties, attributes, and interpretation. Section 3 summarizes the specific applications of the regionalization procedure in the test runs of the TG-B model set planned under Phase 1. It also highlights some important details about ARAM. More-detailed reports that describe the use of ARAM to generate driver data for each sector model are referenced.

1.2 OVERVIEW OF ENERGY-ECONOMIC DRIVER MODULE

Figure 1 presents an overview of the energy-economic driver module in the TG-B emissions model set. This diagram identifies the components of the driver module and maps the flows of data through it. The purpose of the driver module is to generate the requisite inputs for each of the sector emissions models. The components of the driver module are required because the input data for the sector models are generally not provided by energy and economic forecasting models. For this reason, a procedure was devised to systematically generate the regional driver data for input to the sector models. The procedure was divided into three components, as illustrated in Fig. 1: National Adjustment and Disaggregation (NAD), the Argonne Regionalization Activity Module (ARAM), and the Argonne Regional Energy Price Simulator (AREPS). Separate documentation has been prepared to describe each of these submodules; this report describes ARAM. Although each submodule performs a different function or task, each is consistent with the national control totals for each scenario.

*For convenience, the word "regional" is used to refer to either a single state or a multi-state region.

National forecasts of energy and economic activity are supplied to these three submodules. Those data are adjusted, disaggregated, and regionalized to provide the appropriate input data to each sector model. In Fig. 1, the national models supplying these forecasts are identified on the left, the regionalization activities in the middle, and the sector models on the right. Any national (or regional) model or forecast can be used to provide input data to the driver module. The input data for the test runs planned under Phase 1 are detailed below.

The energy and economic driver models indicated in Fig. 1 provide consistent inputs for each energy-economic scenario. Separate but mutually consistent economic models provide macroeconomic, regional activity, and detailed industry projections. These models are part of the Data Resources, Inc., (DRI) system.* The DRI macroeconomic model is a large, simultaneous-equation, econometric model that yields gross national product, personal income, industrial production, employment, housing starts, and financial variables on a national level. The macroeconomic model accepts the Bureau of Census middle projections of population growth as an input. The DRI Regional Information Service (DRI/RIS) model yields state-level projections of population, housing starts, personal income, and employment (by sector) consistent with the national projections. Finally, the DRI Interindustry Service model provides national projections of output by individual industries [i.e., output is given for about 400 four-digit Standard Industrial Classification (SIC) codes].

Key projections of energy prices and consumption are taken from the National Energy Policy Plan (NEPP) prepared by the U.S. Department of Energy (DOE). These NEPP projections are based on the DOE WOIL/FOSSIL2 model, which uses the system dynamics approach. For the Phase 1 test runs, the national projections of energy consumption used by ARAM are taken from the draft April 1985 version of the NEPP (NEPP-85) and include electricity demand by sector, nuclear and renewable electricity generation, cogeneration, and industrial fossil-fuel use. National projections of energy prices by fuel type and sector are also taken from the NEPP-85 but are regionalized in AREPS.

1.3 REGIONALIZATION METHOD INCORPORATED IN ARAM

As indicated, national energy and economic projections must be regionalized for input to the sector emissions models. The proposed regionalization approach incorporated in ARAM has three main steps. The data associated with these steps are illustrated in Fig. 1 by the three boxes with arrows pointing to ARAM. The first two steps refer directly to "driver variables," which are sector emission model inputs appropriately defined to meet the needs of each model. For example, one driver variable for the Industrial Combustion Emissions (ICE) model is consumption of purchased fossil fuel by boilers in the chemical industry (SIC 28). The third step comprises a projection of regional economic activity; these projections are used to derive shift-share factors that

---

*The DRI modeling system was selected for the Phase 1 test runs after an examination of alternative systems.

are applied to the base-year driver variables. An overview of this three-step approach is provided below. A more detailed description is given in Sec. 2.

In the first step of the regionalization procedure, a base-year (1980) data file is assembled. This data file contains estimates of each driver variable at the state level (or regional level as appropriate). This step ensures that model inputs are initialized to their historic values. In Step 2, a national projection for each driver variable is constructed, a process that may require adjustments or disaggregation of the national forecasts from the NEPP or DRI.* For example, nonpurchased fuel use must be subtracted from total industrial fuel use reported in the NEPP to construct the appropriate inputs for the ICE model. Thus national trends in the driver variables, such as the rate of energy conservation, are subsequently reflected at the regional level.‡ In Step 3, shifts in projected regional shares are calculated on the basis of differential regional growth in population or employment by industry group. The DRI/RIS model provides these forecasts of regional growth. Only the economic activities captured in these regional growth forecasts are reflected in the shift-share component of ARAM. Because of this, there are situations (discussed in Sec. 1.4) where ARAM is more or less applicable as a regionalization scheme.

1.4 CONDITIONS FOR APPLICATION OF ARAM

Although the three-step regionalization approach embedded in ARAM has four desirable properties (identified in Sec. 2.1) and efficiently generates the required sector model input data, it is more applicable in some circumstances than in others. The ARAM approach is most applicable where:

- Base-year values continue to be a significant proportion of projected values during the forecast time horizon.
- Sectors are aggregated at a level that allows particular industries to enter or exit the market, as they do on a statistical basis, without upsetting the integrity of the sectoral aggregation.
- The driver variable being forecasted correlates well with the exogenously forecasted activity variable.
- A systematic and consistent regionalization approach is desired for assessment purposes.

These conditions illustrate that ARAM can be used in most regionalization applications. There are cases, however, where the ARAM approach would be less applicable. These

---

*These adjustment or disaggregation activities take place in the NAD module shown in Fig. 1.

‡See Sec. 2.4 for a further explanation of this property inherent in ARAM.

cases are principally when emerging (i.e., previously unidentified or heretofore unknown) industries or activities are excluded from the national control forecast, the regional activity forecast, or the base-year file.

The ARAM approach relies extensively on the national control forecast. That is, the accuracy and comprehensiveness of the sector forecasts at the national level are of critical importance to the regional forecast of sector activity. For instance, if an emerging industry (e.g., biotechnology) is not incorporated in the national forecast, it will not be reflected in the ARAM's regional forecast, even though it may be an important industry to particular regions.

When an industry or activity is excluded from the national forecast, a discrepancy could arise in the regional allocation of energy or economic activity. Such a discrepancy would arise because the ARAM approach relies on the relative growth in regional activities for each sector forecast. If an industry or activity was omitted from the control forecast, the regional projection produced by ARAM would not reflect the true allocation of economic or energy activity. Furthermore, any forecast of an emerging industry or activity is reliable only for a period of 2-3 years from the start of that industry. Predictions of activity for industries that do not exist in the base year are very speculative beyond 3 years into the future, and as such represent exogenous forecasts that are based primarily on judgment. For this reason (which is a principal condition in the application of ARAM), exogenous adjustments are not included as part of ARAM and this Phase 1 exercise.

If an industry or activity does not exist in a region in the base year, it will not be reflected in the future when employing the ARAM regionalization approach. Two examples are

- Renewable energy. This energy type is negligible or nonexistent in all states except California in the base year (1980). As such, any forecast of growth will be allocated to California unless otherwise adjusted.

- Industry. Manufacturing and nonmanufacturing industry groups exist in most states, although in particular states some industry groups do not exist. Therefore, ARAM would not project any activity in states without those industries.

In both of these examples, exogenous information could be used to adjust the future distribution of activity. However, the ARAM approach would no longer be as systematic if many exogenous adjustments were implemented. Consequently, the ARAM regionalization methodology is most applicable when regional activity and base-year input data do not require exogenous adjustment.

2 ARGONNE REGIONALIZATION ACTIVITY MODULE

The Argonne Regionalization Activity Module was developed to derive regional and state projections of energy and economic variables from national forecasts. As such, ARAM bridges the gap between national and regional modeling. It transforms the control values for energy and economic variables that are produced by national forecasting models into the regional projections required as input to the sector models of the TG-B emissions model set.

The properties of the regionalization procedure incorporated in ARAM are described in Sec. 2.1, its mathematical derivation in Sec. 2.2, and interpretations in Secs. 2.3-2.5. The regionalization algorithm design is based on a modified shift-share approach, an approach commonly used in regional economic studies and models. Beginning with base-year (1980) values for driver variables by state, and taking into account national growth, regional shift factors are applied to forecast the required driver variables. The regional shift factors reflect the change in regional shares of an economic activity variable over time. The shift factors are computed from a forecast of economic activity variables related to the driver variables, such as employment in the corresponding industry.

2.1 PROPERTIES OF THE REGIONALIZATION METHODOLOGY

The four properties specified below define the regionalization formula:

1. A national-level projection of each driver variable is provided, incorporating national trends like the rate of energy conservation.

2. Regional quantities must sum to the national projection (i.e., the additivity property holds).

3. The regional driver variable in the base year must equal its historic value. The regional distribution of the driver variable is assembled in a data base for the base year. (The base year for the test runs conducted in Phase 1 is 1980.)*

4. The regional growth index of a driver variable is proportional to the growth index in a related activity variable from the DRI/RIS model forecast. Equivalently, the difference in growth rates in a driver variable between two regions is equal to the difference in growth rates in the regional activity variable.

---

*Where base-year 1980 data for the nation and the regions come from different sources, possible discrepancies are resolved to achieve additivity in the base year.

These properties can be more useful when expressed in symbolic notation:

1. $E_n(t)$, the national driver projection, is given for t = 1980, 1985, ..., 2030 (t = 0 denotes t = 1980).*

2. $\sum_{r=1}^{m} E_r(t) = E_n(t)$ where $E_r(t)$ is the regional driver variable for m regions.

3. $E_r(0)$ is given for all regions.

4. $G_r(t) = Z(t)A_r(t)$ where:

   $G_r(t) = E_r(t)/E_r(0)$, or the regional growth index of the driver variable,

   $A_r(t) = L_r(t)/L_r(0)$, or the regional growth index of the regional activity forecast $L_r(t)$, and

   $Z(t)$ = proportionality factor (uniform for all regions), which may change with time to capture national trends such as increased energy productivity (i.e., less energy use per employee).

Property 4 may be rewritten as follows:

$$\frac{G_1(t)}{G_2(t)} = \frac{A_1(t)}{A_2(t)}$$

This expression implies, as stated above, that the differential growth rate between regions for the driver variable is the same as the differential growth rate for the regional activity variable. It also implies that, if region 1 grows faster than region 2 (i.e., $A_1 > A_2$), region 1 will gain in its share of the driver quantity relative to region 2.

---

*When annual data are needed, and the national projections are provided in 5-year time increments, linear interpolation is used.

2.2 REGIONALIZATION FORMULA

The four properties listed above imply a unique formula for regionalization. This formula, which is the basic regionalization equation in ARAM, can be stated as

$$E_r(t) = E_n(t) \frac{E_r(0)\ A_r(t)}{\sum_{r=1}^{m} E_r(0)\ A_r(t)} \tag{1}$$

where:

$E_r(t)$ = regional projection of the desired driver variable,

$E_n(t)$ = national projection (control value) of the desired driver variable,

$E_r(0)$ = base-year (1980) value of the desired driver variable, and

$A_r(t)$ = regional growth index of a related activity variable.

Consequently, the projected regional share, as computed by ARAM (Eq. 1), must be as follows:*

$$\frac{E_r(t)}{E_n(t)} = \frac{E_r(0)\ A_r(t)}{\sum_{r=1}^{m} E_r(0)\ A_r(t)} \tag{2}$$

Consequently, Eq. 2 is the regional share form of the basic ARAM formula (Eq. 1).

Equation 1 satisfies the the additivity property, because summing the ARAM equation over all regions yields $E_n(t)$. In addition, for $t = 0$, the right-hand side of the equation equals the base-year value $E_r(0)$.‡ Finally, Property 4 holds, since the growth index for the regional driver variable $E_r(t)/E_r(0)$ is proportional to the regional activity growth index $A_r(t)$ in Eq. 1 with the (region-independent) proportionality factor $Z(t)$ equal to

$$Z(t) = \frac{E_n(t)}{\sum E_r(0)\ A_r(t)} \tag{3}$$

---

*From Property 4, the regional projection is given by $E_r(t) = E_r(0)Z(t)A_r(t)$. Summing $E_r(t)$ yields $E_n(t)$. Taking the ratio of $E_r(t)$ to $E_n(t)$ and cancelling $Z(t)$ yields the ARAM regional share formula (Eq. 2).

‡$A_r(0) = 1.0$ and $E_n(0)$ equals $\sum E_r(0)$ in the denominator.

Here Z(t) is a scale factor, uniform for all regions, used to reconcile different aggregate growth rates between driver and activity variables.

2.3 INTERPRETATION OF THE REGIONALIZATION FORMULA

The ARAM formula (Eq. 1) is usefully interpreted in terms of growth rates. As an example, suppose that ARAM predicts a growth of 25% in a regional driver variable between 1980 and 2000.* Then the regional growth index $G_r(t)$ is calculated by ARAM as 1.25. A forecast in year t = 2000 is simply

$$E_r(2000) = G_r(2000)\ E_r(0) = 1.25\ E_r(0).$$

The ARAM formula can be written explicitly in the growth index form as follows:

$$G_r(t) = G_n(t)\ \frac{A_r(t)}{\sum_{r=1}^{m} W_r\ A_r(t)} \tag{4}$$

where:

$G_n(t) = E_n(t)/E_n(0)$, or the growth index in the national projection of the driver variable, and

$W_r = E_r(0)/E_n(0)$, or the regional share of the driver quantity in the base year.

This expression is obtained by dividing both sides of Eq. 1 by $E_r(0)$ and dividing the numerator and denominator on the right-hand side of the equation by $E_n(0)$. Since the weights W_r sum to one, the denominator in the ARAM formula can be interpreted as the weighted average of the regional activity growth indexes. Thus

$$\bar{A}(t) = \sum W_r\ A_r(t) \tag{5}$$

where $\bar{A}(t)$ is the weighted-average index of regional activity.

---

*There is no need to compute the average annual growth rate R_t over the t-year forecast period, but, if this is desired, the formula is

$$R_t = [G_r(t)]^{1/t} - 1$$

or, in the example, $(1.25)^{1/20} - 1 = 0.011$ or 1.1% average annual growth. In general, growth rates from year to year may change, and hence the average annual growth rate over the entire period will also change with the length of the period t.

Equation 4 can then be rewritten as

$$G_r(t) = G_n(t)\,\frac{A_r(t)}{\bar{A}(t)} \tag{6}$$

The interpretation of the regionalization approach specified above is as follows: The regional index for the driver variable $G_r(t)$ grows at the national rate with an adjustment for specific regional differences. The regional adjustment factor is the relative growth in the related regional activity variable (e.g., employment) compared to the average growth in this activity over all regions, i.e., $A_r(t)/\bar{A}(t)$. The weighted average index of regional activity $[\bar{A}(t)]$ need not exactly equal the national growth in this activity, because the regional weights W_r are based on the driver variable $E_r(0)/E_n(0)$ and not the activity variable $L_r(0)/L_n(0)$.

In addition to the above explanation of Eqs. 5 and 6, the interpretation of Z(t) can be made more precise. From Property 4 of the regionalization formula, Z(t) must equal $G_n(t)/\bar{A}(t)$, which is the national-level growth index in the driver variable compared to the weighted-average activity growth index. Hence, Z(t) is interpreted as the national trend in "energy productivity change" or the "change in energy per unit of activity."

2.4 REGIONAL INTENSITY INTERPRETATION

Residential electricity use per person or boiler fossil-fuel use per worker are examples of energy intensity. In general, the term "energy intensity" denotes the ratio of the driver variable to the related activity variable. For example, in the industrial volatile organic compounds (VOC) model, this ratio is uncontrolled emissions per worker in an associated industry.

Energy intensity can be defined at both the national and regional level. Empirically, energy intensities vary widely among states and regions. This variation arises from differences in industrial composition and other differences. The ARAM formula preserves these differences in energy intensity.* More precisely, the ARAM formula implies that energy intensities increase or decrease at the same rate in each region.

To illustrate this, denote energy intensity by X so that

$$X = E/L \tag{7}$$

---

*More elaborate formulas than ARAM have been derived that allow regional energy-intensities to move over time toward the national average. However, there may be little empirical basis for determining the amount of convergence or the rate of adjustment.

where E is the driver quantity and L is the activity variable. (Recall that the growth index for L is denoted by A.) After some manipulation of Eq. 4, the ARAM regionalization formula can be specified in the energy-intensity form

$$\frac{X_r(t)}{X_r(0)} = \frac{X_n(t)}{X_n(0)} \left[\sum_r \frac{X_r(0)}{X_n(0)} S_r(t) \right]^{-1} \tag{8}$$

where $S_r(t)$, defined as $L_r(t)/L_n(t)$, is the regional share of employment (or other activity variable). Nothing on the right-hand side of this equation is unique to a specific region. Therefore, the percentage change in energy intensity as measured by $X_r(t)/X_r(0)$ is equal for all regions. This uniform regional percentage change in energy intensity may be larger or smaller than the national percentage change in energy intensity as measured by $X_n(t)/X_n(0)$.* Hence, with the ARAM formula, energy intensity changes by a uniform percentage across regions, under the condition that regional driver quantities add to the national projection.

2.5 SUMMARY INTERPRETATION

Suppose that the driver variable is some form of industrial energy use (e.g., electricity or boiler fuel) and the related activity variable is employment in a particular industry. Three ingredients are needed to project regional use of industrial energy according to the ARAM formula: (1) base-year (1980) values for industrial energy use in each region, (2) a national forecast of industrial energy use, and (3) a regional forecast of employment in the subject industry.

The ARAM formula allocates total national industrial energy use to component regions in some forecast year, say t = 2000. The percentage change in energy use per worker (comparing the year 2000 with 1980) is the same for every region. Finally, given the formulation of ARAM, optimal conditions for its application exist (see Sec. 1.4).

---

*For example, suppose that large employment regions were relatively more energy-intensive than small regions. That is, suppose that employment shares S_r tend to be larger in those regions where base-year regional energy intensity $X_r(0)$ is greater than national energy intensity $X_n(0)$. Then the summation in Eq. 8 will be greater than one, and hence regional energy intensities will increase less (or decrease more) than those of nation.

3 APPLICATIONS

Applications of the ARAM formula to the driver data required by each sector model in the TG-B emissions model set are described in detail in Refs. 1-4. The main applications, however, are summarized below: end-use electricity demand (Sec. 3.1), industrial boiler fuel demand (Sec. 3.2), uncontrolled industrial VOC emissions (Sec. 3.3), and industrial production indexes (Sec. 3.4). Besides these applications, the ARAM formula can be (and has been) applied to many other regionalization efforts, as long as the requisite input data are provided.*

The summary descriptions in Sec. 3.1-3.4 document how ARAM generates regional driver data for input to each of the TG-B sector models. In Sec. 3.1, a numerical example further illustrates how ARAM generates input required by a particular sector model. The example selected is end-use electricity demand, which is used to drive the Advanced Utility Simulation Model (AUSM) in the test runs planned for Phase 1 of the TG-B program.

3.1 END-USE ELECTRICITY DEMAND

3.1.1 Summary Description

Electricity demand, which drives AUSM, is regionalized to the state level for four end-use sectors: residential, commercial, manufacturing, and other industrial. Greater accuracy and reliability are attained by regionalizing each sector forecast individually. National projections of electricity demand are based on the NEPP-85. The regional activity variables for residential and commercial electricity demand are forecasts of state population and commercial employment, respectively. The regional activity variables for manufacturing and other industrial electricity demand are manufacturing employment and other industrial employment, respectively. As stated earlier, forecasts of regional activity are supplied by the DRI/RIS model.

---

*In the Phase 1 test runs of the TG-B emissions model set, ARAM was also used to derive a special set of input data for AUSM.[5] To derive these input data, end-use electricity demand was adjusted for cogeneration, hydroelectric generation, other renewable generation, and nuclear generation. ARAM was used to generate the regionalized data for each of these adjustments. However, the use of ARAM for providing state-level hydroelectric generation was slightly different than the other applications. One difference was that the control totals were independent regional projections instead of national forecasts. Another difference was that the shift-share factor in the ARAM formula was electricity demand rather than employment or population.

3.1.2 Example Application of ARAM

To illustrate the use of ARAM, this section provides a numerical example. The example demonstrates how ARAM uses a national projection to generate a regional forecast of end-use electricity demand. Figure 2 presents a national forecast of electricity demand by the four major end-use sectors. This projection for the years 1980 through 2030 represents the reference case scenario in the NEPP. For illustrative purposes, the remainder of this example applies the ARAM regionalization methodology to manufacturing electricity demand in the year 2000.

Figure 3 shows an application of the ARAM methodology to manufacturing electricity demand in the year 2000. Several aspects are apparent. First, the result: a regional projection of manufacturing electricity demand in 2000. Second, the two key factors in the shift-share component: regional distribution of manufacturing electricity demand in the base year (1980) and the activity (i.e., manufacturing employment) growth index in the year 2000 relative to 1980.

In the South Atlantic region, for example, manufacturing employment grew by 15% between 1980 and 2000, as indicated by the growth index of 1.15 shown. Alternatively, manufacturing employment in the Midwest region declined by 1%, as indicated by the growth index of 0.99. Because the activity index is growing faster in the South Atlantic region than in the Midwest, ARAM projects that the South Atlantic region will gain in its share of manufacturing electricity demand relative to the Midwest. That is, the shift-share factor in ARAM depends on the activity index.

This dependence is illustrated intuitively in Fig. 3. Note that in 1980 the two regions (i.e., the South Atlantic and Midwest) had approximately equal shares of

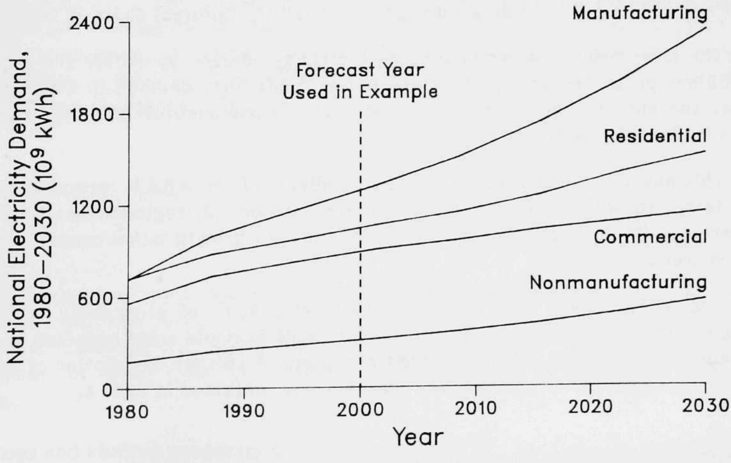

FIGURE 2 Projections of National Electricity Demand by End-Use Sector (Source: NEPP-85)

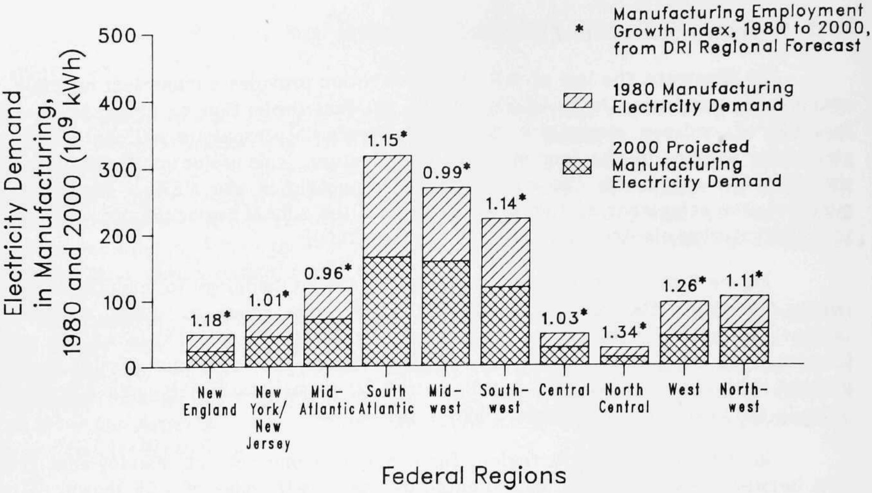

FIGURE 3 ARAM Projections of Regional Electricity Demand in the Manufacturing Sector: 1980 and 2000

manufacturing electricity demand; base-year quantities in both regions were approximately 150 x 10^9 kWh. However, by the year 2000, the projected demand for electricity in the manufacturing sector is significantly greater in the South Atlantic region than the Midwest (total height of bars in Fig. 3). Hence, the South Atlantic share of manufacturing electricity demand is greater than the Midwest share in 2000.

The final results of regionalizing electricity demand by sector and region in the year 2000 are presented in Fig. 4. Here actual electricity demand in 2000 is shown by region as the sum of each sector forecast. The importance of electricity demand by sector in each region is clearly evident.

This example illustrates the three ingredients of the ARAM formula: (1) national control totals to which the regional forecasts balance, (2) regional distribution in the base year, and (3) activity growth index. The activity growth index controls the change in state shares over time.

The AUSM model requires state-level projections of electricity demand. The projections for federal regions given in the previous example were obtained by summing the projections for the states in each federal region. A further description of how ARAM was used to prepare the driver data for the AUSM is contained in Ref. 4.

3.2 INDUSTRIAL BOILER FUEL DEMAND

The Industrial Combustion Emissions (ICE) model requires, as one of its inputs, a state-level projection of boiler fuel demand. This projection is limited to purchased

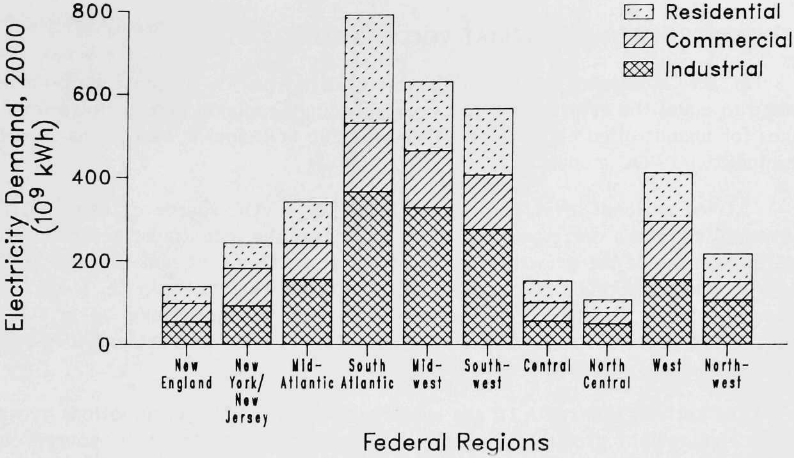

FIGURE 4 ARAM Projections of Regional Electricity Demand Projections by Sector for the Year 2000

fossil fuels consumed in the boilers of seven industry groups.* ARAM is used to prepare these regional projections after several steps are followed to generate an appropriate national control forecast. National projections of industrial fossil fuel use are taken from the NEPP-85. These NEPP projections must be adjusted and disaggregated to produce national forecasts of purchased boiler fuel demand by industry group before they can be used as input to ARAM (see Ref. 6). The industry groups are Standard Industrial Classification (SIC) codes 20, 22, 26, 28, 29, 33, and 99.‡ State employment projections for each industry group, obtained from the DRI/RIS model, are used as the regional activity variables in ARAM. The detailed steps followed to prepare the driver data for the ICE model are contained in Ref. 1.

---

*The fossil fuels considered in the ICE model are purchased natural gas, residual and distillate fuel oil, and coal. Nonpurchased residual fuel oil consumed at refineries is also included. Excluded from consideration are nonpurchased (or interplant transfers of) blast furnace or coke oven gas, refinery off-gas, wood wastes, and hydrocarbon feedstocks.

‡SIC 20, food and kindred products; SIC 22, textile mill products; SIC 26, paper and allied products; SIC 28, chemicals and allied products; SIC 29, petroleum refining and related industries; SIC 33, primary metal industries; and SIC 99, defined here as all other industries.

3.3 UNCONTROLLED INDUSTRIAL VOC EMISSIONS

In the industrial VOC model, the growth rate in uncontrolled emissions is assumed to equal the growth rate in a corresponding economic index. State-level growth indexes for uncontrolled emissions $G_r(t)$ in over 100 VOC source categories must be input to the industrial VOC model.

At the national level, the growth index for a VOC source category $G_n(t)$ equals the growth index in a corresponding economic index; the index must be constructed as a weighted average of the growth rates in economic sectors that make up the VOC source category. These weights (not to be confused with the weights W_r in Eq. 4) are the shares of base-year VOC emissions in the economic sectors that make up a VOC source category. The economic sectors are quite detailed -- projections are provided by the DRI Interindustry Model at the three- or four-digit SIC code level.

The activity indexes $A_r(t)$ are constructed from DRI/RIS projections by weighting regional employment growth in industries corresponding to the VOC source category. Once again, the weights are base-year VOC emission shares. Using these data, ARAM projects state-level growth indexes for uncontrolled VOC emissions by source category. A more detailed discussion of these procedures is found in Ref. 3.

3.4 PHYSICAL INDUSTRIAL PRODUCTION

The driver variable defined for the Industrial Sector Technology Use Model (ISTUM) is a growth index of physical production. In general, the driver inputs to ISTUM are industrial production indexes at the two-digit SIC code level (an exception is steel production). However, output from a two-digit industry group is not homogenous. Different products within an industry group generate different quantities of value added and are produced with different energy intensities (Btu per unit of output). Hence, the desired concept of "physical production" by industry can only be approximated by some form of index number.

Usually the Federal Reserve Board (FRB) index of industrial production is used. For some major energy-intensive industries, however, the FRB index historically has grown much faster than more reasonable measures of physical output. This problem is particularly evident in the chemical industry (SIC 28), but also in pulp and paper (SIC 26) and primary metals (SIC 33). The source of the problem seems to be a trend toward faster relative growth in products with higher value added per unit in the mix of outputs (i.e., a compositional effect). For example, energy-intensive, raw-material processing often grows at a slower rate than the industry as a whole, due to faster growth in downstream, highly fabricated, or specialty products.*

---

*One major reason the FRB index grows faster than a better measure of physical output is that the FRB index weights growth in different products according to their contribution to value added, not their contribution to energy use.

The differences in growth rates between the FRB indexes and better measures of physical output are already incorporated in ISTUM. These differential growth rates have been updated for the Phase 1 test runs.[5] We assume that the differential growth rates between FRB indexes of industrial production and physical output, which are estimated at the national level, can be applied uniformly across regions.

To regionalize physical industrial output with the ARAM formula (for input to ISTUM), appropriate notation must be introduced. At the national level, the growth index in physical output $G_n(t)$ is related to the FRB industrial production index $I_n(t)$ by

$$G_n(t) = B(t)\ I_n(t) \tag{9}$$

where B(t) accounts for the estimated difference in growth rates. [B(t) functions have been estimated in Ref. 5.] We assume that the same differential growth rates apply to the regional level so that

$$G_r(t) = B(t)\ I_r(t) \tag{10}$$

where $I_r(t)$ is a regional index of industrial production. Substituting these relationships (Eqs. 9 and 10) into the growth index form of the ARAM formula (Eq. 4) and cancelling the B(t) from both sides of the equation yields

$$I_r(t) = I_n(t)\ \frac{A_r(t)}{\sum_{r=1}^{m} W_r\ A_r(t)} \tag{11}$$

which is the industrial index form of the ARAM formula.

In applying Eq. 11, $I_n(t)$ is the industrial production index projected by DRI for each industry at the national level. $A_r(t)$ is the employment growth index for that industry by federal region. DRI state-level forecasts of industry employment are aggregated to the ten federal regions to construct $A_r(t)$. The weights W_r are ideally base-year shares of "physical output." However, because of the conceptual problems of defining physical output for heterogeneous industries, base-year regional shares of energy use are used as a proxy.* Energy use is a reasonable proxy, since it is often assumed that energy use, everything else being equal, is proportional to physical output.

Under these assumptions, the ARAM formula can be used to generate regional indexes of industrial production, $I_r(t)$. The ISTUM model applies the growth rate correction to this regional index of industrial production to project growth rate of physical output. This amounts to applying the B(t) function shown in Eq. 10. Reference 2 addresses the procedure in more detail.

---

*For a similar application of energy-weighted output indexes as predictors of energy use, see Ref. 6.

4 SUMMARY

This report describes a method for regionalizing the energy and economic variables used as input to sector models in the TG-B emissions model set. The methodology is first described in general terms, then it is derived and its properties and interpretation presented. The methodology is incorporated in a system called ARAM, the Argonne Regionalization Activity Module.

ARAM links the national and regional modeling activities associated with the Phase 1 test runs of the TG-B emissions model set through its ability to disaggregate national forecasts of energy and economic variables (i.e., as produced by NEPP and DRI) into the regional projections required as input to the sector emissions models. ARAM generates the regional projections by applying a modified shift-share approach. Beginning with base-year (1980) values for driver variables by state, and taking into account national growth, regional shift factors are applied to forecast the required driver variables. The regional shift factors reflect the change in regional shares of an economic activity variable over time. The DRI/RIS model provides the projections from which the regional shift factors are computed.

ARAM can generate the required input data for each sector model in the TG-B emissions model set. The methodology can also be applied to other data variables, given that the data needed by ARAM are available. On the basis of the extensive review undertaken to recommend a regional economic model capable of providing input data to the sector emission models, it can be concluded that ARAM is a unique system.[7] The regionalization procedure contained in ARAM was designed and developed exclusively to provide the sector models with input data. The model review could not identify a modeling system capable of performing this interface function, so the methodology described here was developed to satisfy the data requirements of the sector models and bridge the gap between national and regional models.

REFERENCES

1. South, D.W., M.J. Bragen, and C.M. Macal, *Industrial Combustion Emissions (ICE) Model: Regionalized Projections of Demand for Purchased Industrial Boiler Fuel*, Argonne National Laboratory Report ANL/EES-TM-302 (June 1985).

2. South, D.W., et al., *Industrial Sector Technology Use Model (ISTUM): Regionalized Projections of Industrial Production Indexes*, Argonne National Laboratory Report ANL/EES-TM-330 (July 1985).

3. South, D.W., et al., *Industrial Volatile Organic Compounds (VOC) Model: Regionalized Projections of Uncontrolled VOC Emissions by Source Category*, Argonne National Laboratory Report ANL/EES-TM-306 (June 1985).

4. South, D.W., et al., *Advanced Utility Simulation Model (AUSM): Regionalized Projections of End-Use Electricity Demand*, Argonne National Laboratory Report ANL/EES-TM-300 (June 1985).

5. Boyd, G.A., et al., *Disaggregation of Industrial Fossil Fuel Use in the 1985 National Energy Policy Plan: Methodology and Results*, Argonne National Laboratory Report ANL/EES-TM-303 (June 1985).

6. Marlay, R.C., *Trends in Industrial Use of Energy*, Science, pp. 1277-1283 (Dec. 14, 1984).

7. South, D.W., J.F. McDonald, and W.H. Oakland, Argonne National Laboratory, unpublished information (June 1985).

APPENDIX:
RELATIONSHIP TO THE LITERATURE

by

John Marinelli and Donald Hanson

The regionalization method used in ARAM has been described previously in the emissions inventory literature. For example, in Ref. A.1, Walden disaggregates state-level estimates to county-level estimates with an identical method. First, an unadjusted estimate for each county was prepared. These estimates were simply scaled so that they summed to the state total. That is, the unadjusted estimates served as shares that were used to allocate some control total.

The ARAM procedure is described mathematically by Eq. 1 of the main text, reproduced below:

$$E_r(t) = E_n(t) \frac{E_r(0)\ A_r(t)}{\sum_{r=1}^{m} E_r(0)\ A_r(t)} \tag{1}$$

where unadjusted regional forecasts, $E_r(0)\ A_r(t)$, are used to construct shares to allocate the total, $E_n(t)$, to regions. These shares in Eq. 1 will generally vary over time because they are dependent on the DRI activity forecasts.

In a casual way, the above may be reminiscent of shift-share analysis.* However, it is important to make the critical distinction between activity and driver variables. In shift-share analysis, the focus is on forecasting a regional *activity* variable. With ARAM, the focus is on forecasting a regional *driver* variable, such as energy use. The ARAM algorithm regionalizes national driver forecasts in a way consistent with historical driver values and both regional and national DRI activity

---

*Distinctions are often made between the uses of shift-share analysis for descriptive purposes and forecasting purposes. The former are considered pragmatically useful and the latter technically controversial. Recent literature surveys of shift-share analysis may be found in Refs. 2 and 3. Reference 3 concentrates specifically on issues relevant to the use of shift share as a tool for forecasting. Further discussions of several refinements, reformulations, and extensions of shift-share analysis have appeared in Refs. 4 and 5.

In traditional shift-share analysis, an activity variable (e.g., employment level or growth rate) is the center of concern. Using shift-share analysis to forecast an activity variable is controversial due to the alleged temporal instability of the so-called competitive component investigated by Brown (Ref. 6), whose results were examined by Gerking and Barrington (Ref. 7) and others.

forecasts. The DRI activity forecasts include behavioral considerations in their theoretical bases.

A regional econometric model that explicitly forecasts driver variables may be ideal in some abstract setting, but money and time considerations preclude this option. Rather, a shift-share-like formula is used to convert regional activity forecasts into regional driver forecasts in the way specified previously. Note that the regional driver forecasts are arrived at via a combination of top down and bottom up procedures. Traditional shift-share analysis is essentially a top down procedure.

To further understand ARAM, recall its assumption that for a given sector i, energy intensity (X = E/L, where X = energy intensity, E = driver variable, and L = activity variable such as employment) grows at a rate g_i^X that is equal for states r and s. That is,

$$g_{si}^{x} = g_{ri}^{x} \tag{A.1}$$

Hence

$$g_{si}^{E} - g_{si}^{L} = g_{ri}^{E} - g_{ri}^{L} \tag{A.2}$$

This expression implies that

$$g_{si}^{E} - g_{ri}^{E} = g_{si}^{L} - g_{ri}^{L} \tag{A.3}$$

Consequently, the state growth differentials for driver and activity variables are equal to the state growth differentials for activity variables forecasted by the DRI regional model.

Thus ARAM essentially adjusts technical coefficients, that is, the ratio of energy use to activity, using information on activity from the DRI regional model. In a commonly used form of shift-share analysis, the state growth rate differential for the activity variable is a constant (see the right-hand side of Eq. A.3). One may question this constancy and assert that such factors as agglomeration economies and factor supply scarcities may tend to influence such a term (Refs. 4 and 8). However, as indicated above, regional activity growth is forecasted by the DRI regional model in our conceptual framework, and hence our method need not suffer from the weaknesses of shift share in forecasting activity variables.

In terms of modifying or generalizing the ARAM formula, one may want to allow the energy intensity growth rate to vary over states. For example, the growth rate in energy intensity in a given state may depend on the level of energy intensity in that state. An extreme case would be where we have regression toward the mean. In this case, energy intensities among states, although not initially equal, would eventually become identical. Intermediate cases where some long-run differences in energy intensities remain would be interesting yet difficult to implement.

REFERENCES FOR APPENDIX

1. *Methodologies for Countywide Estimation of Coal, Gas, and Organic Solvent Consumption*, EPA-450/3-75-06, prepared by Walden Research Division of Abcor for U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Research Triangle Park, N.C. (Dec. 1975).

2. Dawson, J., *Shift Share Analysis: A Bibliographic Review of Technique and Applications*, Vance Bibliographics, Monticello, Ill. (1982).

3. Stevens, B.H., and C.L. Moore, *A Critical Review of the Literature on Shift Share as a Forecasting Technique*, Journal of Regional Science, *20*:419-38 (1980).

4. Berzeg, K., *A Note on Statistical Approaches to Shift Share Analysis*, Journal of Regional Science, *24*:277-85 (1984).

5. Haynes, K.C., and Z.B. Machund, *Considerations in Extending Shift Share Analysis: Note*, Growth and Change, *18*:67-78 (1987).

6. Brown, M.J., *Shift and Share Projections of Regional Economic Growth: An Empirical Test*, Journal of Regional Science, *9*:1-18 (1969).

7. Gerking, S., and J. Barrington, *Are Regional Share Effects Constant over Time?*, Journal of Regional Science, *21*:163-74 (1981).

8. Borts, G., and J.L. Stein, *Economic Growth in a Free Market*, Columbia University Press, New York City (1964).

X